Modeling the NASA Baseline and SVS-Equipped Approach and Landing Scenarios in D-OMAR

Stephen Deutsch and Richard Pew
BBN Technologies
Cambridge, Massachusetts

Foyle, D.C., Goodman, A. and Hooey, B.L. (2003). Modeling the NASA baseline and SVS-equipped approach and landing scenarios in D-OMAR. In D.C. Foyle, A. Goodman & B.L. Hooey (Eds.), *Proceedings of the 2003 Conference on Human Performance Modeling of Approach and Landing with Augmented Displays*, (NASA/CP-2003-212267), 143-164. Moffett Field, CA: NASA.

This page intentionally left blank.

1. Introduction

Human performance modeling (HPM) is a technology that has the potential to help address flight deck and air traffic controller (ATC) workplace design issues and support aircrew procedure evaluation. It can be used early in the workplace development process before prototypes or full-scale simulations are available or later in the design cycle to support on-going development. Unfortunately, the technology is not yet sufficiently mature that it can be applied easily and routinely. However, it is reasonable to undertake exploratory studies to evaluate aircrew procedures for employing new flight deck technologies and to evaluate their role in promoting aviation safety. At the same time, these investigations will improve the architectures and software tools that are available to support the human performance modeling process.

Under the aegis of the Aviation Safety Program, NASA Ames Research Center (ARC) has initiated a program element concerned with modeling human performance. The element's goals include advancing the state-of-the-art in human performance modeling and demonstrating their potential payoffs in the commercial aircraft design and aircrew procedure development by contributing to improved aviation safety. Within the framework of modeling human performance, achieving a better understanding of the sources of human error and identifying procedures for error mitigation are a particular focus. This element is a multi-year, multi-contractor effort that is just completing its second year.

For the second year effort, NASA asked that the modeling teams examine approach and landing operations, and compare aircrew performance for baseline and synthetic vision system (SVS) equipped flight decks. BBN's effort has focused on the further development of captain and first officer human performance models and the approach and landing procedures that employ baseline and SVS-equipped flight deck configurations. Human performance models for the approach, tower, and ground controllers have been developed to interact with the aircrew models to faithfully represent the airspace. The human performance modeling effort paralleled and profited from the NASA part-task simulation studies (Goodman, Foyle, Hooey, & Wilson, 2003) that collected pilot performance data in baseline and SVS-equipped flight deck operations at the Santa Barbara Municipal Airport (SBA).

The modeling effort was accomplished using BBN's Distributed Operator Model Architecture[1] (D-OMAR) to represent the behaviors of the aircrews and air traffic controllers. D-OMAR was also used to model the aircraft and their flight decks, the ATC workplaces, and the essential features of the Santa Barbara Municipal Airport and the local airspace. Our goal for this year was to produce models that could appropriately represent successful approach and landing performance similar to that exhibited by the pilot subjects in a subset of the recently conducted NASA SVS part-task simulation experiments. To date, the D-OMAR models are successfully executing five of the part-task scenarios.

In this document, Section 2 traces the development of the human performance models, the aircrew and air traffic controller procedures that they execute, and the strategy for building the emulation of the part-task scenarios. Section 3 then provides a discussion of the findings that resulted from the modeling effort, Section 4 outlines the implications of the findings, and

[1] The D-OMAR software with user manual is available as OpenSource at http://omar.bbn.com.

Section 5 highlights the lessons learned in the process of developing the human performance models to implement the scenarios.

2. Modeling the Approach and Landing Scenarios

2.1 NASA Support for the Modeling Effort

NASA Ames made available several important resources to facilitate this year's modeling effort. The part-task simulation dictated an RNAV rather than an ILS approach as modeled for last year's O'Hare scenarios (Deutsch & Pew, 2001: Deutsch & Pew, 2002). To assist the modeling teams in the transition, a cognitive task analysis (Keller & Leiden, 2002a) provided a detailed description of aircrew and air traffic controller procedures for an RNAV approach. The document included detailed information on how flight deck systems are used to support an RNAV approach. An addendum to the document (Keller & Leiden, 2002b) extended the task analysis to cover the aircrew's use of an SVS during the approach and landing. These documents provided much of the information necessary to construct the D-OMAR goal, sub-goal, and procedure framework that represents the standard aircrew and air traffic controller procedures for an RNAV approach and landing.

Our modeling effort relied heavily on the documentation of and data from the NASA baseline and SVS part-task scenario trials. Goodman *et al.* (2003) provided a detailed description of the simulated flight deck, the design for the ten scenario trials, and a description of the scenario trial data for three subjects. Trial data included simulation output for each run, eye tracker data, and video (with audio) recordings based on an eye-tracker camera and a room-view camera. The eye tracker data included both fixation sequence and dwell sequence data files. The data was made available by trial with summary statistics available by phase of flight. A summary fixation sequence spreadsheet was assembled to provide duration percentage for viewing each instrument across all trials and all subjects. The spreadsheet made it possible to compare and contrast individual subject eye tracking behaviors across the trials. In summary, these data provided a comprehensive view of aircrew behaviors essential to modeling the baseline and SVS-assisted approach and landing trials.

The BBN team also participated in and profited from an SVS/SWAP information-sharing workshop held at NASA Langley late in 2001. The workshop provided a number of presentations on the development of the NASA SVS system as well as a review of the flight test experiments in using the system. The SVS section of the NASA Aviation Safety Program (AvSP) web pages provided access to a number of publications that included the Concept of Operations for Commercial and Business Aircraft Synthetic Vision Systems (Williams, Waller, Koelling, Burdette, Doyle, Capron, Barry, & Gifford, 2001).

2.2 Strategy for Developing the Scenarios

The HPM-SVS part-task simulation study (Goodman *et al*., 2003) included ten scenarios that selectively covered three variables of interest: display configuration, visibility, and approach events. There were two display configurations: a baseline flight deck consisting mainly of a primary flight display (PFD), a horizontal situation indicator (HSI), and a mode control panel (MCP), and a second configuration in which the baseline configuration was augmented with an SVS display. Landing gear, flaps, speed brakes, throttle settings, and map scale for the HSI were set by the first officer using a separate display panel.

Visibility conditions included visual meteorological condition (VMC) and instrument meteorological condition (IMC). The light haze in VMC allowed visual flight rules. Under IMC, there was a *reported* 800-foot ceiling. Approach events included a nominal approach, a "late reassignment" requiring a sidestep to a parallel runway, a missed approach requiring a go-around, and a "terrain mismatch" in which the SVS was found to be misaligned as the aircraft emerged from the cloud cover. The "go-around" scenarios took two forms: in the VMC case, traffic on the runway made the go-around necessary. In the IMC case, in spite of the reported 800-foot ceiling, the aircraft was still in the cloud cover at the 650-foot decision height, hence the captain was unable to acquire the runway making the go-around necessary.

As modelers, we were asked to focus on the "late reassignment" scenarios using the baseline and SVS-equipped flight decks. Our approach was to first develop the basic aircrew procedures to support the RNAV approach and landing using the baseline flight deck under VMC. We then added weather to represent IMC. In the modeled IMC condition, there was an *actual* cloud ceiling at 800 feet; hence, the breakout from the cloud cover occurred at 150 feet above the designated decision height.

For the "late reassignment" condition, we developed a scenario with two closely spaced aircraft on the approach to SBA 33L. As the lead aircraft landed, it blew a tire and temporarily held on the active runway. This set up a situation requiring the ATC to request that the following aircraft side step to the parallel runway 33R. Lastly, we added an SVS to the flight deck and extended the aircrew procedures to include the use of the SVS for the IMC scenario and the "late runway reassignment" scenario.

At this point, the modeled aircrews successfully execute the approach, landing, and taxi procedures for five of the NASA part-task simulation scenarios:

- Scenario 1 – nominal approach using the baseline configuration in VMC
- Scenario 2 – late reassignment approach using the baseline configuration in VMC
- Scenario 4 – nominal approach using the baseline configuration in IMC
- Scenario 7 – nominal approach using the SVS configuration in IMC
- Scenario 8 – late reassignment approach using the SVS configuration in IMC

2.3 Human Performance Models for the Aircrew and the Air Traffic Controllers

In the scenarios, the captain and first officer must work together to safely execute the approach-and-landing procedures. In responding to a series of controller directives, actions must be prioritized, appropriate aircrew communications must be generated to coordinate the execution of these actions, and interrupts in the form of further directions from the controllers must be handled. The interrupts are not unexpected, but rather meet expectations consistent with the local air traffic and weather conditions. Reactive behaviors are determined within the framework of the aircrew's active goals and procedures. In meeting their responsibilities, the captain and first officer have a significant number of tasks in process, each of which requires a coordinated mix of perceptual, cognitive, and motor skills. The scenarios create situations in which the response to demands must be carefully prioritized to achieve acceptable performance.

Each aircraft is populated by human performance models for the captain and first officer. The aircrew models are extensions of last year's models that executed the O'Hare ILS approach, landing, and taxi (Deutsch & Pew, 2001: Deutsch & Pew, 2002). The new RNAV procedures that they employ at SBA are based on the Keller and Leiden (2002a; 2002b) cognitive task analysis. Consistent with our long-term goal of examining error mitigation for two person aircrews, the procedures that our models execute closely follow the Keller and Leiden real world task analysis. Hence, there are departures from the exact procedures as tailored for the part-task simulation trials where the captain was the experiment subject and the first officer was a surrogate minimally supporting the captain as necessary.

In the same spirit, the modeled approach and landing scenarios follow the standard progression in the airspace transiting from approach controller to tower controller to ground controller and terminating as the aircraft completes the landing and taxies to its assigned concourse. In the modeled scenarios, each controller is represented by a human performance model. The scenarios as modeled more closely follow real world operations than was possible in the part-task simulation.

The aircraft model includes the instruments and controls necessary for the crew to execute the required approach and landing scenarios. The model more closely emulates the actual 757 flight deck rather than part-task simulation flight deck. The principal instruments include the PFD, the HSI, and the SVS. The PFD includes annunciators for LNAV and VNAV. Controls include the MCP, switches for the autopilot, and levers for the throttles, flaps, landing gear, and speed brakes. The central instrument panel includes lights providing landing gear status.

The aircrew makes use of the approach plate for SBA runway 33L for information related to the RNAV approach. The information used by the captain to brief the go-around contingency at the beginning of the approach is derived from the approach plate. The approach plate is subsequently used by the aircrew as a reference to track the sequence of fixes for the approach path and as the source for the required descent altitude for each approach leg. As they depart the active runway, the first officer uses the SBA airport diagram to support taxi operations leading to the concourse.

Voice communication between the captain and first officer is used to coordinate the execution of approach and landing procedures. Party-line radio communication is modeled with the aircrew resetting radio frequencies as they move from one controller to the next. Careful attention has been paid to the fine details of interleaving of aircrew and air traffic controller conversations and to handling air traffic controller interruptions to aircrew conversations.

As each scenario begins, the approach controller clears NASA186 (and NASA277 for the two aircraft scenarios) for the approach to SBA 33L and provides information on VMC or IMC depending on the scenario. Using information from the approach controller and the approach plate for runway 33L, the captain continues the approach by reviewing the runway and weather information with the first officer. Go-around procedures are briefed, based on the approach plate information. For the RNAV approach, the captain asks that MCP modes LNAV and VNAV be set and checks the PFD annunciator mode lights.

The aircrew then focuses on navigation as the aircraft proceeds along the flight management computer (FMC) flight path from one fix to the next. As they approach each fix, the captain calls for a new MCP altitude setting for the next leg. The aircrew monitors the aircraft's heading and altitude changes (information derived principally from the HSI) as the aircraft transitions onto the leg to the next fix. They continue to monitor the heading until the new desired heading is fully established. They monitor altitude to assure that they hold at the designated target altitude. As the approach progresses, the captain calls for a series of flap settings consistent with their speed and position along the approach path. The handoff from the approach controller to the tower takes place along the leg to GOLET.

The aircrew contacts the tower and is immediately given the clearance to land. The captain asks for the final flap setting, that the landing gear be lowered, and that the speed brakes be armed. At this point, the captain asks for execution of the landing checklist, which is then acted on with the first officer. As the aircraft's descent continues, the first officer monitors the aircraft's altitude and makes call outs at 1000 feet above field level (afl), as they approach decision height, and at 100 feet afl. The captain is responsible for the out-the-window sighting of the runway and making the decision to land. Under VMC, the captain can readily acquire the runway out the window well before decision height. For the SVS-equipped, IMC-condition scenario, the captain can use the SVS to acquire the runway before they break out of the cloud cover, but must still acquire the runway out the window to make the decision to land. The captain anticipates the break out from the cloud cover at 800 feet and has adequate time to acquire the runway before the 650-foot decision height. The captain must take manual control of the aircraft to preempt the preprogrammed go-around and further manage the landing.

Figure 1 provides a D-OMAR plan view of two aircraft on their approach to SBA 33L in the "late reassignment" scenario. At the point presented in the plan view, the first aircraft has blown a tire on landing causing it to hold temporarily on the active runway. The top panel on the right in Figure 1 records the conversations among the approach and tower controllers and the two aircraft. NASA277's communication with the tower controller related to the blown tire shows up in the recorded dialog. The tower controller addresses the situation by asking the second aircraft, NASA186, to sidestep to runway 33R.

The lower panel on the right records details of the conversation on the flight deck of the following aircraft, NASA186. Diane, the NASA186 captain, tells her first officer that she will accept the request to sidestep to 33R. The first officer communicates the captain's acceptance of the sidestep request to the tower controller. The first officer's response completes the sidestep transaction with the tower controller related in the upper display panel. The captain then announces her decision to proceed with the landing. As they land, the captain manages speed brake and reverse thrust settings, while the first officer calls out current ground speed. Following the landing, NASA186 notifies the tower as the aircraft leaves the active runway, contacts ground controller and receives directions to taxi to the it's assigned concourse.

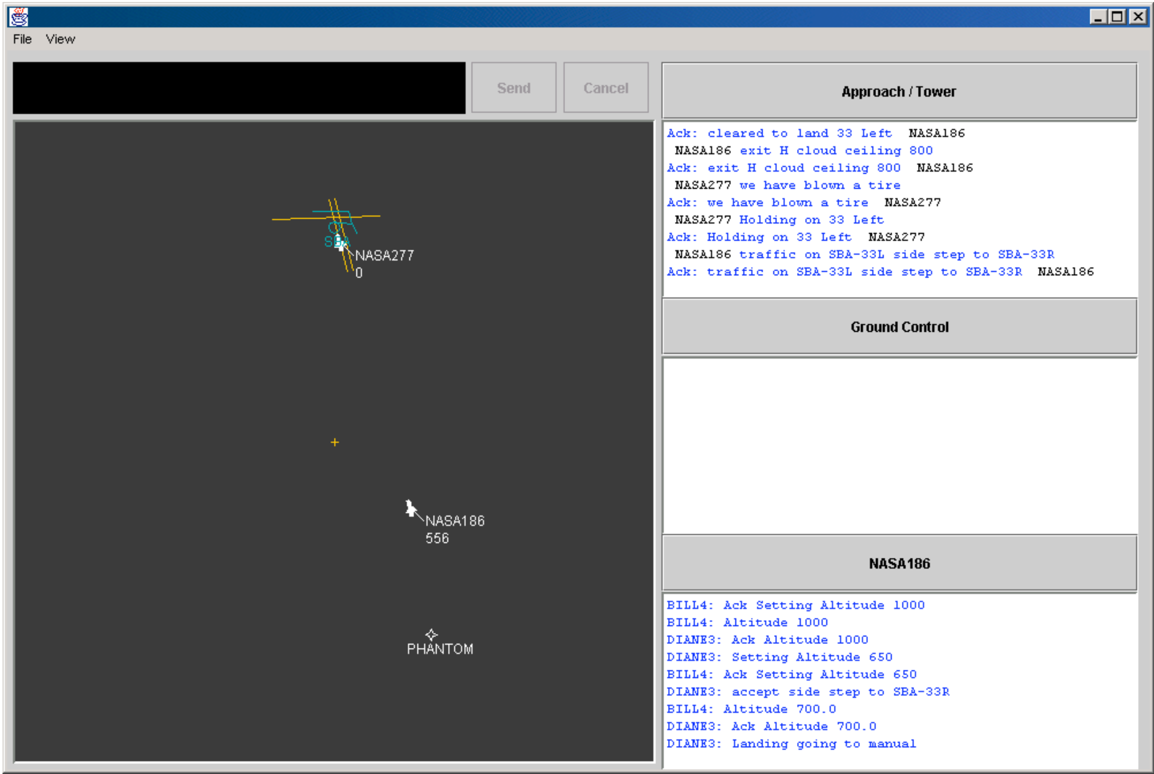

Figure 1. Screen View from the Late Reassignment Scenario.

2.4 Assessing Model Behaviors in D-OMAR

D-OMAR simulation tools provide explicit measures of model behaviors. They are essential in the assessment of model performance just as they are essential for managing the complexity required to create the models. Aircrew behaviors are frequently multitask behaviors—the more, or occasionally less, successful integration of the demands of several ongoing procedures. Each of these procedures is made up of several steps that require the coordinated interaction of several human functional capabilities (e.g., the maintenance of a conversation, the coordination of hand-eye actions to set a selector). The execution of a checklist interrupted by an ATC communication is at once a common occurrence and a challenging event sequence to faithfully model. Multiple levels of visibility into model behaviors are essential both to develop the scenarios and to assess model performance.

D-OMAR graphical display tools each provide a unique view into model behaviors. A plan view, as illustrated in Figure 1, allows an observer to monitor the progress of the aircraft along its flight path. The plan view display has recently been supplemented by a similar HSI-like display (see Figure 4). A Gantt chart display (see Figure 3) provides detailed information on goals and procedures as executed by the captain and first officer. An event timeline (not illustrated) provides detailed insight into the behaviors of the publish-subscribe protocol used to coordinate procedure execution. Lastly, a detailed event trace is recorded for each simulation run with key events displayed in the trace pane of the simulation control panel (see Figure 2) as the simulation progresses.

Some of the evaluation tools operate concurrently with the simulation; others are used once a simulation run has completed or after the simulation is paused. The plan views and the simulation trace operate concurrently with the simulation, the task and event timelines are available once the simulation has been paused. An event recording system is used to capture the data to support most of the evaluation tool presentations. Several event types are basic elements of the D-OMAR simulator, others are more specialized and created to address the data capture requirements of a particular domain or scenario. Procedure execution is the basic element driving agent behaviors; hence, events are recorded that identify the agent executing the procedure, the beginning and end times for each procedure, the success or failure of the procedure, the procedure's priority, the name of its parent procedure, and any time periods during which the procedure was interrupted.

Several event types have been created to track the performance of the human performance models for the aircrew and air traffic controllers. One of the event types records flight deck actions taken by each aircrew member (and workplace actions for the air traffic controllers); for example, the setting of an MCP selector for altitude or the movement of the lever to establish a particular flap setting. Since in-person conversations on the flight deck and party-line conversations with the air traffic controllers are so important, conversation events have been defined to record these conversations. The aircrew and air traffic controller conversations presented in the right hand panels of Figure 1 are generated by "after methods" on the event recording process.

The event types for flight deck actions and conversations are key elements providing data for the on-line trace. They record and print the actions taken by the aircrew and the air traffic controllers. As such, they represent the outcomes of the execution of the goals and procedures whose development was based on the cognitive task analysis for the RNAV approach and landing (Keller & Leiden, 2002a).

The simulation control panel (Figure 2) provides an interface for the user to select and manage the execution of a scenario, and an on-line trace of selected scenario events. The "Scenario" line in the panel provides for the selection of the scenario to be executed. The "Initialize," "Run," and "Pause" buttons enable the user to control the flow of scenario execution. The D-OMAR simulator is capable of both real-time and fast-time operation. In fast-time mode, the simulator is very efficient taking about 40 seconds to complete 1150 seconds of real-time for the two aircraft "late reassignment" scenarios.

The sample panel shown in Figure 2 includes a short section of the trace from the nominal VMC approach scenario, the SBA-RNAV-VMC-SCENARIO. At this point in the scenario, the captain has just asked that the landing gear be lowered, the flaps be set to 25 degrees, and the speed brakes be armed. As recorded in the trace, the first officer attends to each request and executes each requested action, first checking current setting and then adjusting the setting as necessary. Part way through the sequence of requests for actions on the part of the first officer, the captain adjusts the MCP IAS/mach speed selector and consistent with established procedures, announces the change as it is made.

The slider at the side of the Figure 2 provides a hint at the size of the trace for the nominal VMC scenario. The trace is lengthy, but manageable in size, in large part, because the printing for most event types is turned off. For example, procedure event data that is essential to the task timeline described below, would if printed, make it far more difficult to isolate the important events currently presented. The ability to capture a complete set of scenario event data, yet tailor the trace content is important in making the trace a useful analysis tool.

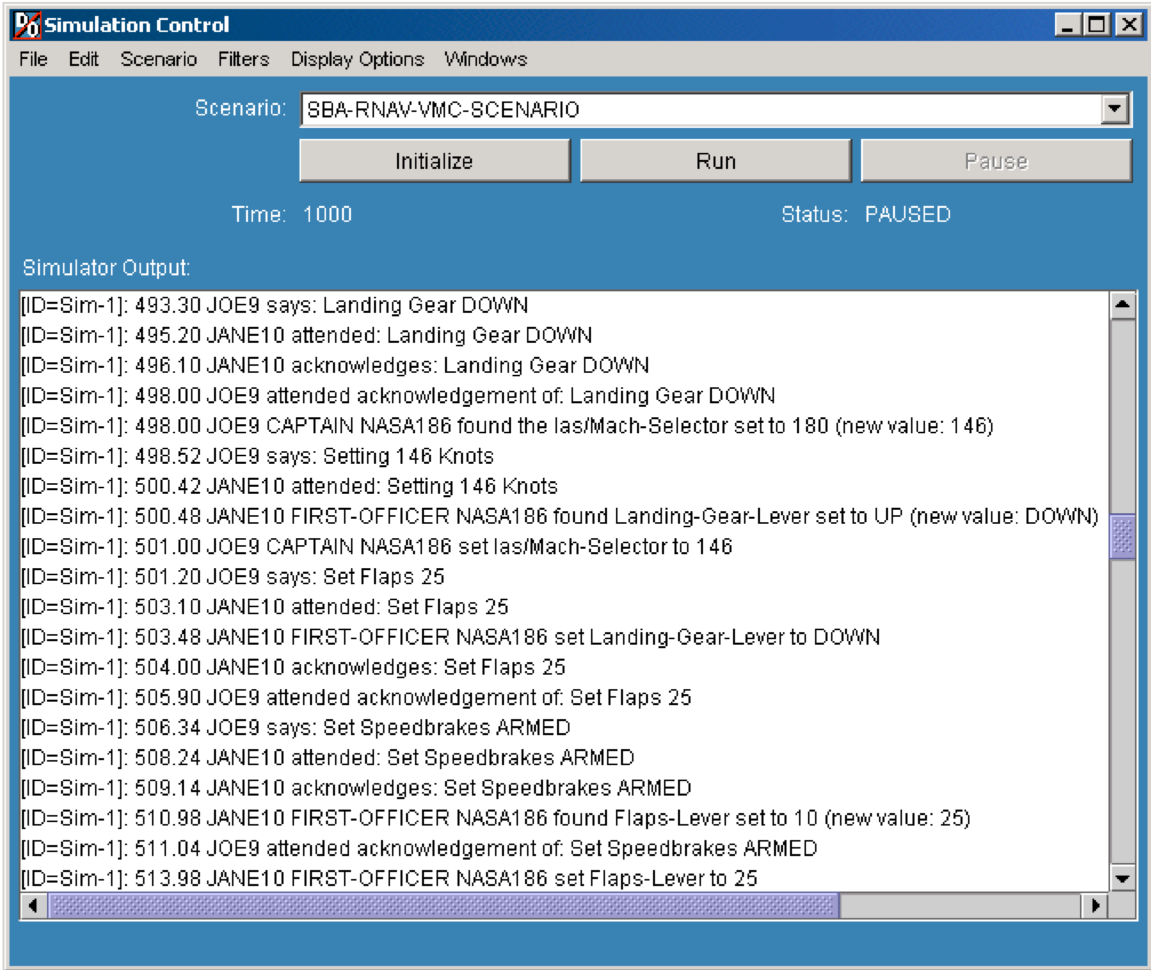

Figure 2. Simulation Control Panel and On-line Trace.

The on-line trace has been specialized to provide insight into critical actions taken by aircrew and air traffic controller models in the scenarios. The actions are the products of the execution of clusters of a captain's or first officer's goals and procedures that each has duration in time. The task timeline display, a Gantt-style display is used to provide insight into how an agent's goals and procedures play out in time to generate these actions. Figure 3 is an example of the timeline display from the baseline VMC scenario. The slider at the right of the figure is quite small indicating that the captain has quite a large set of on-going goals and procedures beyond those currently visible in the display. Some of these goals and procedures represent preparedness to respond to anticipated events, some are currently active in the timeframe covered by the display. The selected timeframe for the display, identified in the display's bottom panel, is fifty seconds, from 550 seconds into the scenario to 600 seconds in the scenario. The aircraft is approaching GOLET and about to transition to the leg to PHANTOM on the approach to SBA 33L.

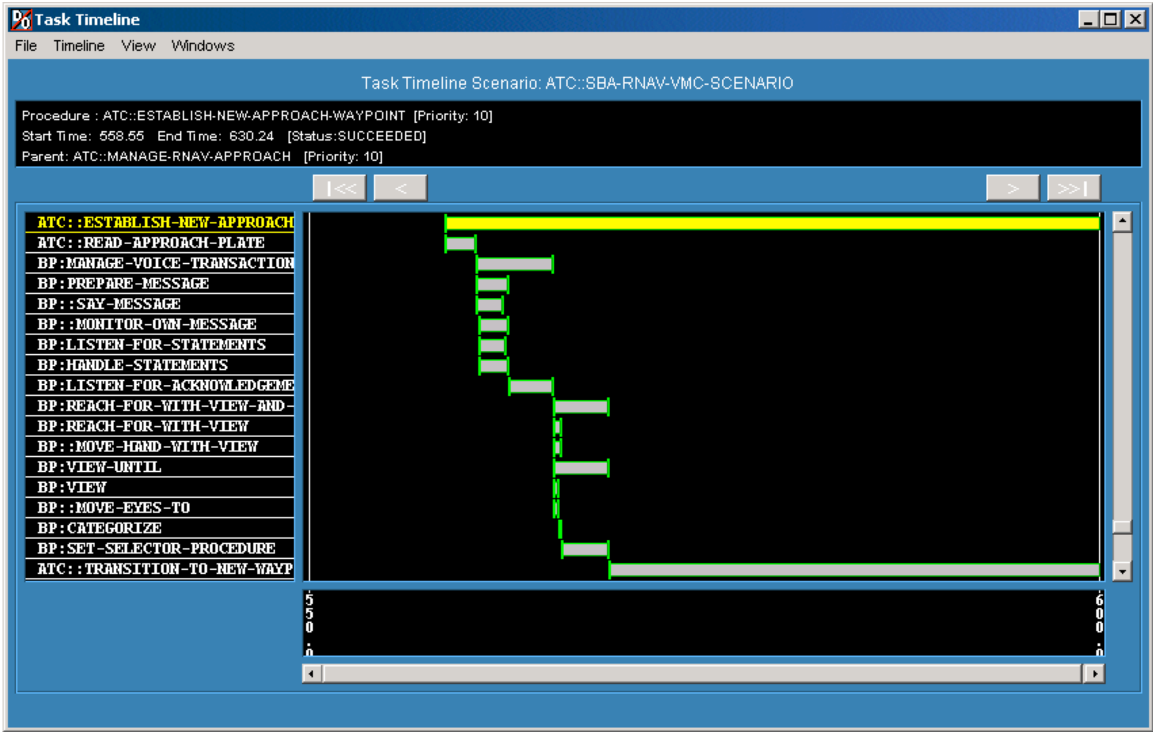

Figure 3. Task Timeline Display.

The procedures that appear in the timeline in Figure 3 are a subset of the captain's procedures as s/he prepares for and then monitors the transition to PHANTOM, the next leg of the approach. Each line of the display represents one of the captain's goals or procedures. The goal or procedure's name appears in the panel at the left, the bar represents the duration of the procedure's execution. Time periods for which a procedure is interrupted (there were none for these procedures) are indicated within the bar representing the procedure's duration. The mouse (not shown), over the first procedure, causes the top panel to be filled with that procedure's primary attributes: the procedure's name and priority, its start time, its end time if it has completed, its current status, and the name of its parent goal or procedure.

In the procedures shown in Figure 3, the captain is concerned with the basic actions for establishing the aircraft on the new leg. The captain first reads the approach plate to verify the altitude for the next waypoint (executed as a single procedure). The captain then tells the first officer that s/he is setting the MCP with the new desired altitude (the next seven lines of the display cover the verbal transaction: speaking the message and listening to the acknowledgement). The captain then sets the MCP IAS/mach speed selector (a coordinated hand-eye activity accomplished by the next eight lines of the display). In the last line shown in the timeline display, the captain is concerned with monitoring the aircraft's transition to the new heading and altitude. Information on current heading is derived from the on-going scan of the HSI display. Information on current altitude is derived from the PFD and the SVS for the scenarios in which it is available. The monitoring procedure obtains this information by "subscribing" to scan procedures that "publish" this information. The publish-subscribe protocol provides a basic mechanism to coordinate procedure execution and to move information between procedures. It is more fully described in the following section.

2.5 Describing the D-OMAR Human Performance Models

The development of human performance models in D-OMAR (Deutsch & Pew, 2002; Deutsch & Pew, 2001; Deutsch, 1998; Deutsch & Adams, 1995) has been based on research in cognitive neuroscience, cognitive science, experimental psychology, and recent cross-disciplinary work in the theory of consciousness. As with most human performance models, the complexity of the models makes it difficult to provide a description that is both brief and complete. There is a theoretical framework that underlies the architecture, a broad range of individual human functional capabilities that must be represented, there are complex interactions among these capabilities, that taken together, generate the observed behaviors, and there are practical compromises that inevitably must be made in producing a working model. To provide additional insight into the D-OMAR human performance model behaviors used in the NASA SBA approach and landing scenarios, it might be useful to briefly examine a few central features of the model: how multiple task behaviors are modeled, the role of vision in supporting the model's multiple task behaviors, and the modeling of working memory.

2.5.1 An Aside on Model Architecture

For convenience within this report, we have spoken of the D-OMAR model, but to be more accurate, that reference needs clarification. D-OMAR itself is simply a general-purpose discrete event simulator. It has been tailored specifically to provide a software framework in which to explore alternate architectures for human performance modeling. The D-OMAR representation languages, a frame language, a rule language, and a procedural language[2] provide the basis for constructing the alternate architectures. The particular models employed for this NASA research task are a further development within an architecture for human performance modeling that has evolved over a number of years.

Most human performance models (e.g., ACT-R (Anderson & Lebiere, 1998), SOAR (Laird, Newell, & Rosenbaum, 1987), EPIC (Meyer & Kieras, 1997), MIDAS (Corker & Smith, 1993)) are implementations of a particular cognitive architecture. D-OMAR, rather than being a particular cognitive architecture, is a simulation framework in which to experiment with and evolve architectures for human performance models. It has been used, in this case, to implement a particular architecture that has evolved to address the NASA task. D-OMAR has readily been used to explore variations on elements of the current architecture and to implement an entirely different architecture. This level of flexibility in model architecture seems essential to the effort to improve the capabilities of human performance models.

2.5.2 Multiple Task Management

One of the principle areas of research in the development of D-OMAR has been in the area of modeling human multitask behaviors. In developing D-OMAR, we have sought to provide a computational framework in which to assemble functional capabilities that operate in parallel, subject to appropriate constraints, and that taken together exhibit the multiple task behaviors of human operators—aircrews and air traffic controllers. The desired behaviors have a combination of proactive and reactive components. That is, the operators have an agenda that they are pursuing, but must also respond to events as they occur. Consequently, within the proactive agenda, there may be newly motivated tasks for which on-going tasks must be deferred. The bounds on what can be accomplished concurrently take several forms. A

[2] Detailed information about the representation languages is available in the documentation for the Lisp version of D-OMAR at the web site http://omar.bbn.com.

typical behavior may be to set aside a flight deck conversation in order to respond to an ATC communication, while at another level, two competing tasks may each require the eyes to guide a manual operation. In the first instance, it is a matter of protocol, in the second, contention for a physical resource.

The core of a D-OMAR model is a network of procedures whose signal-driven activation varies in response to events that are proactively channeled to achieve the operator's goals. From a bottom up perspective, there is an assembly of individual perceptual, cognitive, and motor capabilities that are recruited as procedures to address current goals and sub-goals. Neumann's (1987) functional view of attention, and the localization of mental operations in the brain, as put forward by Posner, Petersen, Fox, and Raichle (1988) are important contributions supporting this approach to modeling human behaviors. Taken together, they point to the functional components in task execution as taking place at particular local brain centers with the coordinated operation of several such centers being required to accomplish any given task. The form that the coordination might take is of particular importance in developing a model of behaviors. A publish-subscribe protocol provides the signal-driven activation needed to coordinate the actions of the various perceptual, cognitive, and motor centers acting in support of the completion of the task. The publish-subscribe protocol also serves to move information among the functional centers.

From a top down perspective, the things that person knows how to do, basic person skills (e.g., coordinated hand-eye actions to set a selector) and domain specific skills (e.g., making the decision to land), are represented as goals, sub-goals, and procedures. Active goals represent the operator's proactive agenda for managing his or her tasks. The goals typically activate a series of sub-goals and procedures. The goals and sub-goals represent the objectives of the actions to be taken; the procedures are the implementation of the actions to achieve the goals and sub-goals. The procedures each may include decision points to address variations in the local situation. Hence, the operator's overall agenda is implemented by the network of procedures established by the goal hierarchy and linked by the publish-subscribe protocol. A subset of the procedures are active, most are in a wait-state—they represent the potential downstream actions of the operator's current actions and his or her ability to cope with a changing world.

Within this framework, process (Edelman, 1987; 1989) has a preeminent role. Basic person skills and domain specific skills encompass far more than simple perceptual or motor skills, they include the highly refined *cognitive* skills that are the mark of significant human expertise (Logan, 1988; Bargh & Chartrand, 1999). Taken together, a model's goals and procedures, the capabilities of the model to perform in a human-like manner, are a major component of the model's long-term memory.

2.5.3 Vision as a Component within Multi-task Behaviors

Demands on the pilot's visual system are varied and complex. A broad range of these visual capabilities is included in the models. Some of a pilot's actions are purely visual. There are basic processes that take in information from the major flight deck instruments and the view out the windscreen. The viewing of a flight deck instrument can be a generic guidance or navigation status check, or the monitoring of an instrument for a specific target value. The view out the windscreen can be to assure that there are no traffic conflicts (i.e., seeing *nothing* can be the desired outcome), acquiring a specific object (e.g., sighting runway 33L on the approach to support the decision to land), or tracking the aircraft's alignment with runway

33R while executing the sidestep maneuver. Acquiring a specific value occurs at brief instant in time, in tracking the runway, the viewing may extend over a fair period of time with breaks to address other visual tasks.

For some actions, the visual component plays a supporting role. The execution of flight deck operations require coordinated hand-eye actions to set switches, adjust selector settings, and reposition control levers.

Reading and more broadly, the interpretation of graphical information are further important visual tasks. The approach plate and airport diagram are used as sources of information to support approach, landing, and taxi operations. Some of the information from these sources is purely textual; some involves the interpretation of annotated graphical information. Each of these visual operations is an important component of model behavior.

Vision plays a central role in the modeling of a pilot's multitask capabilities. The pilot's procedures for each of these visual activities has a priority associated with it. Within this framework, the vision system is a resource and the tasks that require its capabilities compete based on their assigned priority. As modeled, the scans of the windscreen and the individual flight deck instruments, the PFD, the HSI, and the SVS when it is present, are modeled as separate procedures, all with the same priority. The background scan pattern thus produced, moves smoothly to the windscreen and from instrument to instrument. The interval between scans for each are adjusted for each instrument as appropriate as the approach and landing progresses (e.g., the scan of the HSI becomes more frequent at flight path waypoints). At particular points in the approach, instruments may be dropped from the scan pattern (e.g., the scan of the HSI after the decision to land). Clark (1999) describes similar variations in purpose directed saccade patterns as first reported by Yarbus (1967). To date, it has not been necessary to adjust the priorities for the basic scan procedures.

The action to scan out the windscreen for the runway or to support coordinated hand-eye actions to set a selector or control lever operate slightly differently. These actions are examples of a decision to take an action *now* and hence, are invoked at a priority higher than the background scan procedures. The elected action takes place immediately and once the action has completed, the background scan procedures resume. Actions with a visual component that extend over time operate with another slight variation. Visually tracking the runway to support the sidestep maneuver operates similarly with respect to priority, but to accommodate the extended time duration it allows intervals at which background visual procedures can intervene.

Information obtained by the pilot models in their scan of the HSI and the SVS is representative of the information derived from the major flight deck instruments. The HSI is a rich information source used by the pilots in tracking their progress along the flight path in each of the scenarios. Some of the information is immediately symbolic: the aircraft's heading, the distance to the next waypoint, and the display scaling. Some is readily determined in symbolic form once a little graphical interpretation has been accomplished: the name of the next waypoint. Some is geometrically interpreted: that the changing heading is converging on the desired heading for the new flight plan leg. Lastly, the pilots recognize the current waypoint as the last one before landing and use this information in deciding when to terminate their scan of the HSI. Figure 4 provides a screen shot of the HSI display as NASA186 is traversing the leg to GOLET. Each of the basic information items can be seen in the figure. The form of the display is based on the HSI display described in the RNAV cognitive task analysis (Keller & Leiden, 2002a).

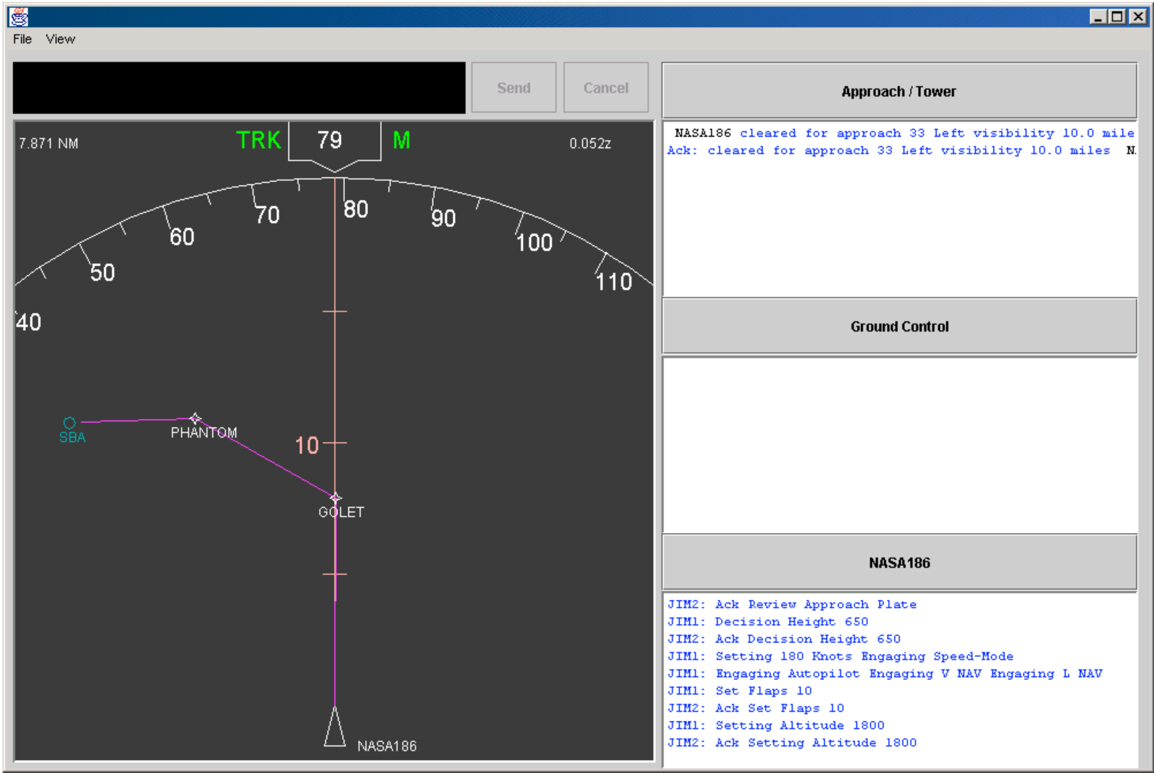

Figure 4. Horizontal Situation Indicator.

The scan of the SVS is particularly important to this research effort. What we now have in place, while sufficient to address the current scenarios, will need to be improved as we address scenarios that are more challenging and as we pursue the identification and mitigation of aircrew errors. For the present, the pilots readily derive heading, speed, altitude, and altitude rate as numeric quantities from the SVS much as they do from the PFD. As the airport comes into the display's field of view, the pilot's view of the runway is based on the distance to the runway. The SVS-equipped scenarios modeled to date are in IMC conditions with an 800-foot cloud ceiling. The pilots are able to track the runway using the SVS before the aircraft breaks out of the cloud cover.

2.5.4 A Distributed Model of Working Memory

On the flight deck, each aircrew member will typically have several tasks at various stages of completion. In pursuing the completion of each ongoing task, there will usually be several goals and procedures concurrently active. As the aircraft proceeds along the flight leg to PHANTOM in IMC with an 800-foot cloud ceiling, at least two of the captain's tasks are concerned with current altitude. For our example, let us say that the altimeter is currently reading 1021 feet afl. The more immediate task will be concerned with monitoring the descent to the target altitude of 1000 feet for the flight leg to PHANTOM. The second task concerned with tracking the current altitude, has as it goal, the decision to land to be made as the aircraft descends through the decision height of 650 feet. There is other work in process, but these two tasks are sufficient to suggest how the altimeter reading is processed in the model.

For all scenarios, an altitude reading is available from the PFD. For those scenarios that include an SVS, the SVS provides a second source for an altitude reading. The captain and the first officer each periodically scan the PDF (and separately, the captain scans the SVS when present). In our example, we will follow the captain's reading of the PFD. The scans each provide, at minimum, the heading, altitude, and speed in numeric form. In the vernacular of the model, upon completion of the reading of the PFD the value of "1021" for the altitude is "published"—that is, the labeled value is part of an attitude scan message that is the product of "scanning" the PFD (or the SVS).

In our example, at least two of the captain's procedures "subscribe" to attitude scan messages, the first concerned with monitoring the descent to the target altitude of 1000 feet for the current flight leg to PHANTOM, and the second preparing to make the decision to land. The active scan procedures for both the PFD and the SVS publish attitude messages, which in effect, "wake" the procedures that are subscribed to the message type. The output of a proactive visual process triggers cognitive processes requiring that output to facilitate their next actions.

From the captain's perspective, the value "1021" is processed in a unique manner in each procedure. For the task of monitoring the descent to the target altitude, "1021" immediately becomes "I need to attend carefully to the aircraft's altitude over the next few seconds to assure that we level off at 1000 feet." For the decision to land task, "1021" becomes "I've got a little time, but I should be breaking out of the cloud cover shortly and I then need to acquire to the runway to support the decision to land." In each case, the published altitude value of "1021" is immediately and significantly reinterpreted by the subscribing procedures. The value "1021" is just an intermediate value in a rapid succession of transformations.

The model's publish-subscribe protocol is designed to mirror the movement of information through the brain's visual centers and on to cognitive centers for further interpretation, action planning, and action execution. The model glosses over the many processing steps in the visual center that produce the symbol "1021," but then more faithfully represents the further processing of "1021" that led to the captain's actions related to assuring that the aircraft levels off at 1000 feet and to preparations to look for the break in the cloud cover and the sighting of runway 33L.

"1021" is certainly a working memory item, but it is simply one local value at one stage in a multi-branching process of transformations and interpretations starting in the visual system and moving through cognitive areas and then to motors areas that drive the resultant actions. Working memory items need a home in the architecture for a human performance model, but they cannot be properly captured in a single box in an architecture diagram. In the D-OMAR model, we posit that working memory, as we have just seen, is widely *distributed* across brain centers. Moreover, we assert that these working memory items do not have a separate existence as database items, but rather are each encapsulated by local processes, the procedures that operate on them at each stage of their migration, transformation, and interpretation.

3 Findings

Aircrew model development and the successful execution of the baseline and SVS-equipped part-task simulation scenarios led to findings related the use of the SVS as a second attitude instrument and findings related to the development of more capable human performance models.

3.1 Successful Scenario Execution

The D-OMAR aircrews, as expected, readily accomplished the five modeled scenarios. For the baseline scenarios in VMC and IMC conditions at SBA, the modeled aircrews successfully executed the approach and landing using RNAV procedures much as the human subjects did in the part-task simulation. The story was much the same for the nominal approach in IMC conditions using the SVS-equipped flight deck. When on the baseline VMC approach and in the SVS-equipped IMC approach, the tower controller requested that the aircrew side step from SBA runway 33L to the closely parallel runway 33R, the captain instructed the first officer to accept the request and then successfully executed the side step maneuver to runway 33R.

Actual performance of the aircrew models was reviewed at several levels of detail either during scenario execution or by reviewing data collected during a simulation run. A time-tagged on-line trace (see Figure 2) tracked the aircrew's conversation on the flight deck as well as the exchanges with the controllers managing the airspace. The trace also tracked flight deck actions taken by the aircrew that followed from this discourse. These traces confirmed that aircrew performance followed the procedures laid out in the Keller and Leiden (2002a) cognitive task analysis. A more detailed view of aircrew performance was reviewed using the Gantt-style display (see Figure 3) of goal and procedure execution for the captain and first officer. This display was used to review and evaluate aircrew performance at the task level by examining the timeline for goal, sub-goal, and procedure execution leading to task completions.

3.2 The SVS as a Second Attitude Display has Workload Implications

The addition of the SVS display to the flight deck augmented the out-the-window view while at the same time providing much of the same functionality as the PFD. In our model, the captain used the SVS to view runway 33L while still in the cloud cover, but reverted to the out-the-window view once the runway came in sight. Interestingly, there were individual differences in the behaviors of the three subjects in the part-task experiment during the flight phase from decision height to landing. While subjects four and five made the expected use of the out-the-window view, subject three relied more heavily on the SVS using the out-the-window view for only five percent of the flight phase.

When the SVS was added to flight deck, the captain, as modeled, included both the SVS and PFD in the scan for aircraft attitude, speed, altitude, and altitude rate information. The basic scan then included the out-the-window view, the PFD, the SVS, and the HSI. One impact of the scan of the two flight deck instruments with an overlap in functionality was that less time was devoted to the HSI display and the navigation function that it supports. Upon further review, the same effect was found in human subject data for the part-task simulation. A finding identified in the modeling process, was further supported by human subject data from the part-task simulation.

In viewing the SVS, there is certainly the sense that some elements of the terrain are more important than others, and that this is situation dependent, changing as the scenario progresses. During a nominal approach to SBA 33L, the target runway, once in range, may well be the most important SVS display feature. However, this would quickly change in the event of a go-around, particularly at an airport such as SBA where there are several hills along the 33L go-around route. Similarly, in the case of the sidestep maneuver to parallel runway 33R, the focus of attention must shift from 33L to the new runway assignment. The temptation is to highlight the salient features in the SVS display and adjust the highlighting to newly salient features as the approach progresses or as the situation changes. The request from the controller to shift the landing to 33R is simply an auditory radio exchange and the pilot assumes manual control of the aircraft to execute the sidestep maneuver. Unfortunately, it is not clear how to include the SVS in the information transfer related to decision to shift the landing to runway 33R. The automation is not a party to the exchange. It receives no information from which to capture the decision to land on 33R and has no basis from which to act to highlight 33R as the new target runway. The decision is conducted behind its back. Data link is a possible path by which to make the privileged information more broadly available on the flight deck. It might then be possible to think about highlighting salient SVS display features.

3.3 Additional Scenario Complexity Can Identify Model Shortcomings

Using a second aircraft in the late reassignment scenario was a small step in adding to the realism of the scenario. It was enough to trigger the need to refine model behaviors in dealing with the conflict that arises when there is air traffic controller communication at the same time that an aircrew has critical actions that they need to coordinate on. The nominal behavior is simply to have the aircrew defer their flight deck conversation when the air traffic controller interrupts a conversation in progress. This basic rule broke down first in the O'Hare scenario when the first officer had to get an okay from the captain before accepting the exit taxiway offered by the approach controller. The captain "spoke through" the on-going air traffic controller exchange to provide the okay to the first officer. In the late reassignment scenario, the lead aircraft was approaching decision height as the tower controller was providing the trailing aircraft with the clearance to land. The lead aircraft's first officer was monitoring the descent to decision height and similarly had to "speak through" the air traffic controller conversation with the trailing aircraft to notify the captain of its approach. Including more complexity in the scenarios teased out flaws in the models and lead to model improvements.

4 Implications

For the five SBA scenarios executed by the D-OMAR aircrew models, the aircrews readily accomplished the approach and landing tasks using the baseline and SVS-equipped flight decks. While aircrew performance readily meet the requirements of the scenarios, the SVS as a second attitude display was found to have implications with respect to aircrew workload. With respect to the models themselves, greater scenario complexity was one means by which to stress model capabilities, identify their shortcomings, and drive the development of more capable models. As model capabilities improve, it will be possible to more effectively address complex aircrew issues in aviation safety challenges.

4.1 Two Attitude Displays or One?

The modeled aircrews and the subjects in the part-task experiments tended to spend less time attending to the HSI display when the SVS was available even in the early phases of the approach where they were principally monitoring their progress along the flight path. The presence of the SVS appeared to reduce the time allocated to the HSI even when the HSI was the information source most relevant to the current flight phase. A simple explanation might be that the aircrews had sufficient time to accomplish their navigation task and were simply using the HSI as required. Nevertheless, it is possible that there is a underlying problem.

On the SVS-equipped flight deck, the aircrews effectively have two attitude instruments. In scanning these two separate instruments, it appears that they may be drawn to spend more time attending to attitude-related information than would be necessary when using a single attitude-instrument configuration. With the dual instrument configuration, the habituated pattern may become an unduly complex two-instrument scan with attitude information derived from the two similar but not identical displays. In situations where time pressure is high, having two instruments from which to obtain required information can impose additional burdens on the pilot. Does the pilot stick with the time consuming habituated scan or try to save time by switching to a single instrument, but less practiced scan? Even electing to consider the options has a cost. Moreover, the extra cognitive effort required to switch to a single instrument scan may negate its potential inherent advantage. Changing a habituated two-instrument scan pattern when change is most difficult is an imposition on the aircrew that should be avoided if possible.

One potential solution is to consider a single attitude display combining PFD and SVS functionality. An SVS that has nominal PFD behavior as a fail-safe mode might, if feasible, be considered and explored. The habituated scan would then be a single instrument scan that readily avoids the choice of scan pattern that the two-instrument configuration can potentially impose, just when choice is most difficult.

4.2 Complex Scenarios Can Drive Necessary Model Improvements

The addition of a second aircraft in the late reassignment scenario proved to be enough to identify a subtle improvement needed in an aircrew model's handling of an ATC communication with another aircraft at a time of high workload—the first officer "spoke through" the ATC communication to let his/her captain know they were approaching decision height. Scenarios that are more complex stress our models, but more importantly drive the development of models that provide better insight into pilot behaviors. Increased workload can negatively impact aircrew performance. Increased workload for human performance models may well promote the development of better models—models that help us to understand resulting changes in aircrew performance, enable us to identify the errors that might be the product of high workload situations, and provide a setting in which to evaluate error mitigation strategies.

5 Lessons Learned

D-OMAR provided excellent support for developing the five initial SBA approach and landing scenarios. The aircrew models that executed the ILS landings at O'Hare for last year's study proved to be readily extendable. The RNAV approach, as detailed in the cognitive task analysis (Keller & Leiden, 2002a), required the implementation of broad range of new goals and procedures, but it was relatively easy to accomplish that within the framework for the human performance models established for the O'Hare scenarios. As the RNAV procedures were developed, they also shared many of the goals and procedures for the ILS approach. As the changes were made, they did at times impact the ILS code and it was necessary to take additional steps to insure the continued operation of the O'Hare scenarios.

In effect, we now have a working RNAV approach and the ILS procedures are more complete.

Constructing the SBA airport model was done using data structures developed for the O'Hare model. Information on Santa Barbara Municipal Airport was obtained primarily from the AIRNAV.COM web page for the airport. The web page included latitude-longitude information for the endpoints of the runways, information on radio frequencies for the controllers, and a link to an airport diagram. Information on runway signage was not available, and was constructed to enable taxi operations similar to those used in the O'Hare scenarios. The AIRNAV.COM pages also provided information on the radio navigation aids for the approach to SBA 33L. The navigation aid information obtained included location latitude and longitude, and radio frequencies. The availability of an airport physical description database would certainly help the process of constructing an airport model. The developers and maintainers of AIRNAV.COM have provided an important first step in this direction.

The NASA part-task simulation data has proven to be a very valuable resource with much still to be learned. In particular, the fixation sequence data files and the eye tracker video tape provide a level of detail in instrument scanning that our models ought to more accurately represent. For the present, the models look at an instrument and read the instrument's data items in a single pass. The eye tracker fixation data videotapes suggest that the pilots selectively and repeatedly scan individual items within a display before moving on to the next display. Clark (1998) reports similar patterns of saccade sites being visited and revisited in the short term. A number of recent studies (Wickens, Xu, Helleberg, Carbonari, & Marsh, 2000; Diez, Boehm-Davis, Holt, Pinney, Hansberger, Schoppek, 2001; Hüttig, Anders, & Tautz, 1999; Anders, 2001) have used eye tracking to examine pilot scan patterns. To the extent that this data is made available, there will be more and more data available to tell us *what* the pilots' scan patterns are. As the developers of pilot models and the theory that underlies our models, we must seek to explain *how* basic human capabilities operate together to yield these scan patterns. Our models will better represent pilot performance to the extent that these behaviors are better explained.

Our long-term goal remains to make use of the understanding of pilot behaviors as represented in human performance models to explore the means to "reduce accidents by mitigating system-wide accident precursors." Aircrews, the pilots in the part-task experiments and the D-OMAR pilot models in the SBA scenarios, readily make appropriate use of the SVS. The scenarios as modeled, add complexity to the scenarios as executed in the part-task scenarios. We would like to build further complexity into the scenarios, refine aircrew procedures for the use of the SVS (possibly eliminating the PFD for some scenarios), and run a series of trials with the aim of probing modeled pilot behaviors for potential benefits as well as errors that might occur on the SVS-equipped flight deck. For human performance shortfalls that lead to errors, we would like to examine approaches to mitigate those errors.

6 References

Anders, G. (2001). Pilot's attention allocation during approach and landing: Eye- and head-tracking research in an A330 full flight simulator. In *Proceedings of the 11th International Symposium on Aviation Psychology*. Columbus, OH.

Anderson, J. R., & Lebiere, C. J. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bargh, J. A, & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54,* 462-479.

Clark, A. (1999). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.

Corker, K. M., & Smith, B. R. (1993). An architecture and model for cognitive engineering simulation analysis: Application to advanced aviation automation. In *Proceedings of the AIAA Computing in Aerospace 9 Conference*. San Diego, CA.

Deutsch, S. E., & Pew, R. W. (2001). *Modeling human error in D-OMAR*. (BBN Report 8328.) Cambridge, MA: BBN Technologies.

Deutsch, S. E., & Pew, R. W. (2002). Modeling human error in a real-world teamwork environment. In *Proceedings of the Twentieth-fourth Annual Meeting of the Cognitive Science Society (pp. 274-279)*. Fairfax, VA.

Deutsch, S. E. (1998). Interdisciplinary foundations for multiple-task human performance modeling in OMAR. In *Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society*. Madison, WI.

Deutsch, S. E., & Adams, M. J. (1995). The operator model architecture and its psychological framework. In *Proceedings of the 6th IFAC Symposium on Man-Machine Systems*. Cambridge, MA.

Diez, M., Boehm-Davis, D. A., Holt, R. W., Pinney, M. E., Hansberger, J. T., & Schoppek, W. (2001). Tracking pilot interactions with flight management systems through eye movements. In *Proceedings of the 11^{th} International Symposium on Aviation Psychology*. Columbus, OH.

Edelman, G. M. (1987). *Neural Darwinism: The theory of neuronal group selection*. New York: Basic Books.

Edelman, G. M. (1989). *The remembered present: A biological theory of consciousness*. New York: Basic Books.

Goodman, A., Hooey, B.L., Foyle, D.C., and Wilson, J.R. (2003). Characterizing visual performance during approach and landing with and without a synthetic vision display: A part-task study. In *Conference Proceedings of the 2003 NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented*

Displays, (David C. Foyle, Allen Goodman & Becky L. Hooey, Eds.). NASA Conference Proceedings NASA/CP-2003-212267.

Huettig, G. Anders, G., & Tautz, A. (1999). Mode awareness in a modern glass cockpit: Attention allocation to mode information. In *Proceedings of the 10th International Symposium on Aviation Psychology*. Columbus, OH.

Keller, J., & Leiden, K. (2002a). *Information to support the human performance modeling of a B757 flight crew during approach and landing: RNAV.*

Keller, J., & Leiden, K. (2002b). *Information to support the human performance modeling of a B757 flight crew during approach and landing: SVS addendum.*

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). SOAR: An architecture for general intelligence. *Artificial Intelligence, 33,* 1-64.

Logan, G. D. (1988). Automaticity, resources, and memory: Theoretical controversies and practical implications. *Human Factors, 30,* 583-598.

Meyer, D. E., & Kieras. D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review, 104,* 3-65.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action*. London: Lawrence Erlbaum.

Posner, M. I., Peterson, S. E., Fox, P. T., & Raichle, M. E. (1988). Localization of cognitive operations in the human brain. *Science, 240,* 1627-1631.

Wickens, C. D., Xu, X., Helleberg, J. R., Carbonari, R., & Marsh, R. (2000). *The allocation of visual attention for aircraft monitoring and avoidance: Baseline measures and implications for freeflight*. Savoy, IL: University of Illinois.

Williams, D., Waller, M., Koelling, J., Burdette, D., Doyle, T., Capron, W., Barry, J., & Gifford, R. (2001). *Concept of operations for commercial and business aircraft synthetic vision systems, Version 1.0.*

Yarbus, A. F. (1967). *Eye movements and vision*. New York: Plenum Press.